

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/627,007  
Source: FWO  
Date Processed by STIC: 12/7/04

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:36

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

```

4 <110> APPLICANT: Gottesman, Susan
5      Storz, Gisela
6      Repoila, Francis
7      Wassarman, Karen
8      Rosenow, Carsten
10 <120> TITLE OF INVENTION: IDENTIFICATION OF NEW SMALL RNAs AND
11      ORFs OF E. COLI AS MEDIATORS OF CELL AND INTERCELL REGULATION
13 <130> FILE REFERENCE: NIH210.001C1
C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/627,007
C--> 15 <141> CURRENT FILING DATE: 2003-07-25
15 <150> PRIOR APPLICATION NUMBER: PCT/US02/03147
16 <151> PRIOR FILING DATE: 2002-01-31
18 <150> PRIOR APPLICATION NUMBER: US 60/266402
19 <151> PRIOR FILING DATE: 2001-02-01
21 <160> NUMBER OF SEQ ID NOS: 161
23 <170> SOFTWARE: FastSEQ for Windows Version 4.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 93
27 <212> TYPE: DNA
28 <213> ORGANISM: E. Coli
30 <400> SEQUENCE: 1
31 gccccttcaa gagctaagcc actgagagtg ccggagataa gcgccggatg gggtagaaac 60
32 ccttaagcct gtgtcgcaca gacttaaggg ttt                                     93
34 <210> SEQ ID NO: 2
35 <211> LENGTH: 86
36 <212> TYPE: DNA
37 <213> ORGANISM: E. Coli
39 <400> SEQUENCE: 2
40 tcgctgaaaa acataaccce taaaatgcta gctgtaccag gaaccacctc cttagcctgt 60
41 gtaatctccc ttacacgggc ttattt                                         86
43 <210> SEQ ID NO: 3
44 <211> LENGTH: 307
45 <212> TYPE: DNA
46 <213> ORGANISM: E. Coli
48 <400> SEQUENCE: 3
49 actgcggccc ttccgccgt ctcgcaaacg ggcgctggct ttaggaaagg atgttccgtg 60
50 ccgtaaatgc aggtgtttca cagcgcttgc tatcgccgca atatcgccag tgggtcgtgc 120
51 gtgatgcggt cttcgcattg accgcacaat gaagatacgg tgcttttgta tcgtacttat 180
52 tgtttctggt gcgctgttaa ccgaggtaaa taataaccgg agtctctccg gcgacaattt 240
53 actggtgggt aacaaccttc agagcagcaa gtaagcccga atgccgccct ttggggcgga 300
54 tattttta                                     307
56 <210> SEQ ID NO: 4
57 <211> LENGTH: 65

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:36

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

```

58 <212> TYPE: DNA
59 <213> ORGANISM: E. Coli
61 <400> SEQUENCE: 4
62 acggcgagc caagatttcc ctggtgttgg cgcagtattc gcgcaccccg gtctagccgg 60
63 ggtca 65
65 <210> SEQ ID NO: 5
66 <211> LENGTH: 92
67 <212> TYPE: DNA
68 <213> ORGANISM: E. Coli
70 <400> SEQUENCE: 5
71 cgcgatcagg aagaccctcg cggagaacct gaaagcacga cattgctcac attgcttcca 60
72 gtattactta gccagccggg tgctggcttt tt 92
74 <210> SEQ ID NO: 6
75 <211> LENGTH: 211
76 <212> TYPE: DNA
77 <213> ORGANISM: E. Coli
79 <400> SEQUENCE: 6
80 aacgagtaga tgctcattcc atctcttatg ttgcgcttag tgccataa actccggaat 60
81 gacgcagagc cgtttacggt gcttatcgtc cactgacaga tgctgcttat gcctcatcag 120
82 acaccatgga cacaacgttg agtgaagcac ccacttggtg tcatacagac ctgttttaac 180
83 gcctgctccg taataagagc aggcgttttt t 211
85 <210> SEQ ID NO: 7
86 <211> LENGTH: 141
87 <212> TYPE: DNA
88 <213> ORGANISM: E. Coli
90 <400> SEQUENCE: 7
91 catcaacacc aaccggaacc tccaccacgt gctcgaatga ggtgtgttga cgtcggggga 60
92 aaccctcctg tgtaccagcg ggatagagag aaagacaaag accggaaaac aaactaaagc 120
93 gcccttggtg cgcttttagt t 141
95 <210> SEQ ID NO: 8
96 <211> LENGTH: 79
97 <212> TYPE: DNA
98 <213> ORGANISM: E. Coli
100 <400> SEQUENCE: 8
101 tgccactgct tttctttgat gtccccattt tgtggagccc atcaaccccg ccatttcggt 60
102 tcaaggttga tgggttttt 79
104 <210> SEQ ID NO: 9
105 <211> LENGTH: 272
106 <212> TYPE: DNA
107 <213> ORGANISM: E. Coli
109 <400> SEQUENCE: 9
110 tgttttaaagc aaaggcgtaa agtagcaccc atagagcgag gacgctaaca ggaacaatga 60
111 ctcaggatga gggtcaggag cgccaggagg cgaagacaga ggattgtcag gaagacaaac 120
112 gtccggagac gtaattaaac ggaaatggaa tcaacacgga ttgttcccta aaggaaaaaac 180
113 aggggtgtgtt ggcggcctgc aaggattgta agaccggtta agggttatga gtcaggaaaa 240
114 aaggcgacag agtaatctgt cgcctttttt ct 272
116 <210> SEQ ID NO: 10
117 <211> LENGTH: 195
118 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:36

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

```

119 <213> ORGANISM: E. Coli
121 <400> SEQUENCE: 10
122 acattgtaaa ccagagttgc gaaggtacaa aaaattaacg ttttagcaat agctatataa 60
123 tatagcctgt gctatatctg tatgtaatgc aatcatccct caaggatcga cgggattagc 120
124 aagtcaggag gtcttatgaa tgagttcaag aggtgtatgc gcgtgttttag tcattctccc 180
125 tttaaagtac ggtaa                                     195
127 <210> SEQ ID NO: 11
128 <211> LENGTH: 82
129 <212> TYPE: DNA
130 <213> ORGANISM: E. Coli
132 <400> SEQUENCE: 11
133 atcccagagg tattgatagg tgaagtcaac ttcgggttga gcacatgaat tacaccagcc 60
134 tgcgcagatg cgcagggtttt tt                                     82
136 <210> SEQ ID NO: 12
137 <211> LENGTH: 92
138 <212> TYPE: DNA
139 <213> ORGANISM: E. Coli
141 <400> SEQUENCE: 12
142 atcccagagg tattgattgg tgagattatt cggtaacgtc tottcgtacc ctgtctcttg 60
143 caccaacctg cgcggatgcg cagggttttt tt                                     92
145 <210> SEQ ID NO: 13
146 <211> LENGTH: 278
147 <212> TYPE: DNA
148 <213> ORGANISM: E. Coli
150 <400> SEQUENCE: 13
151 actataaagt cagcgaagga aatgcttctg gcttttaaca gataaaaaga gaccgaacac 60
152 gattcctgta ttcgggtccag ggaaatggct cttgggagag agccgtgcgc taaaagttag 120
153 cattaatgca ggcttagttg ccttgccctt taagaataga tgacgacgcc aggttttcca 180
154 gtttgcgtag aaaatgggtc ataaaaagcg tgggtggtcat cagctgaaat gttaaaaacc 240
155 gcccggttctg gtgaaagaac tgaggcgggt tttttatt                                     278
157 <210> SEQ ID NO: 14
158 <211> LENGTH: 105
159 <212> TYPE: DNA
160 <213> ORGANISM: E. Coli
162 <400> SEQUENCE: 14
163 agggcaaggc aactaagcct gcattaatgc caacttttag cgcacggctc tctcccaaga 60
164 gccatttccc tggaccgaat acaggaatcg tgttcgggtc ctttt                                     105
166 <210> SEQ ID NO: 15
167 <211> LENGTH: 144
168 <212> TYPE: DNA
169 <213> ORGANISM: E. Coli
171 <400> SEQUENCE: 15
172 agtgagggtt agggagaggt ttccccctcc ccttggtggt cttagtaagc ctggaagcta 60
173 atcactaaga gtatcaccag tatgatgacg tgcttcatca taaccctttc cttattaaaa 120
174 gccctcttct ccgggagagg cttt                                     144
176 <210> SEQ ID NO: 16
177 <211> LENGTH: 137
178 <212> TYPE: DNA
179 <213> ORGANISM: E. Coli

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:36

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

```

181 <400> SEQUENCE: 16
182 agtgagggtta gagcgggggtt tcccccgccc tggtagtctt agtaagcggg gaagcttatg 60
183 actaagagca ccacgatgat gagtagcttc atcatgaccc tttccttatt tatggccccct 120
184 tcctcgggag gggctttt                                     137
186 <210> SEQ ID NO: 17
187 <211> LENGTH: 112
188 <212> TYPE: DNA
189 <213> ORGANISM: E. Coli
191 <400> SEQUENCE: 17
192 aggaacaagg gtaaggagg atttctcccc cctctgattg gctgttaata agctgcgaaa 60
193 cttacgagta acaacacaat cagtatgatg acgagcttca tcataaccct tt      112
195 <210> SEQ ID NO: 18
196 <211> LENGTH: 139
197 <212> TYPE: DNA
198 <213> ORGANISM: E. Coli
200 <400> SEQUENCE: 18
201 cagggcaata tctctcttgc aggtgaatgc aacgtcaagc gatgggcgtt gcgctccata 60
202 ttgtcttact tccttttttg aattactgca tagcacaatt gattcgtagc acgccgactt 120
203 tgatgagtcg gctttttttt                                     139
205 <210> SEQ ID NO: 19
206 <211> LENGTH: 155
207 <212> TYPE: DNA
208 <213> ORGANISM: E. Coli
210 <400> SEQUENCE: 19
211 tagagtaaag gaacaagggt aaggagggat ttctcccccc tctgattggc tgttaataag 60
212 ctgcgaaact tacgagtaac aacacaatca gtatgatgac gagcttcac tcataaccctt 120
213 ccttctgtaa ggcccccttc ttcgggaggg gctttt                                     155
215 <210> SEQ ID NO: 20
216 <211> LENGTH: 128
217 <212> TYPE: DNA
218 <213> ORGANISM: E. Coli
220 <400> SEQUENCE: 20
221 cataggggca atgataaaaag gtggcaaaaa tgaatgtttc cagtagaact gtagtactga 60
222 taaatttctt tgctgctgtt ggtttgttta ctcttatctc tatgagattt ggctgggtta 120
223 tttgatgt                                     128
225 <210> SEQ ID NO: 21
226 <211> LENGTH: 31
227 <212> TYPE: PRT
228 <213> ORGANISM: E. Coli
230 <400> SEQUENCE: 21
231 Met Asn Val Ser Ser Arg Thr Val Val Leu Ile Asn Phe Phe Ala Ala
232 1          5          10          15
233 Val Gly Leu Phe Thr Leu Ile Ser Met Arg Phe Gly Trp Phe Ile
234          20          25          30
237 <210> SEQ ID NO: 22
238 <211> LENGTH: 84
239 <212> TYPE: DNA
240 <213> ORGANISM: E. Coli
242 <400> SEQUENCE: 22

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:36

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

```

243 ataattataa gagaggttgt tatgattgaa cgtgaactgg ggaactggaa agactttatc 60
244 gaagttatgc ttcgtaagta attc                                     84
246 <210> SEQ ID NO: 23
247 <211> LENGTH: 19
248 <212> TYPE: PRT
249 <213> ORGANISM: E. Coli
251 <400> SEQUENCE: 23
252 Met Ile Glu Arg Glu Leu Gly Asn Trp Lys Asp Phe Ile Glu Val Met
253   1             5             10             15
254 Leu Arg Lys
258 <210> SEQ ID NO: 24
259 <211> LENGTH: 180
260 <212> TYPE: DNA
261 <213> ORGANISM: E. Coli
263 <400> SEQUENCE: 24
264 aaaggagacg cttatgtttc gttggggcat catatctctg gttatcgcgt taatcgccgc 60
265 cgcacttggg tttggtggtc tggccggtag cgtcgcaggc gcagctaaaa ttgtctttgt 120
266 cgtcgggatt attctgttcc tggtagagtt gttcatgggc cgaaaacgac cctagatttc 180
269 <210> SEQ ID NO: 25
270 <211> LENGTH: 53
271 <212> TYPE: PRT
272 <213> ORGANISM: E. Coli
274 <400> SEQUENCE: 25
275 Met Phe Arg Trp Gly Ile Ile Phe Leu Val Ile Ala Leu Ile Ala Ala
276   1             5             10             15
277 Ala Leu Gly Phe Gly Gly Leu Ala Gly Thr Ala Ala Gly Ala Ala Lys
278             20             25             30
279 Ile Val Phe Val Val Gly Ile Ile Leu Phe Leu Val Ser Leu Phe Met
280             35             40             45
281 Gly Arg Lys Arg Pro
282             50
285 <210> SEQ ID NO: 26
286 <211> LENGTH: 226
287 <212> TYPE: DNA
288 <213> ORGANISM: E. Coli
290 <400> SEQUENCE: 26
291 atacggagat atcatcatgg gcaaattagg tgaaaacgtt ccgcttctta tcgataaaagc 60
292 cgtagatttc atggcatcaa gccaggcgtt ccgggagtat ctgaaaaaac ttcctccccg 120
293 taacgcgatt ccgtccggaa taccgatga aagcgtgccg ttatatctac aacgtctgga 180
294 gtattatcgt cggttttatt ggccgaagca ggtagagggg cagtaa             226
296 <210> SEQ ID NO: 27
297 <211> LENGTH: 69
298 <212> TYPE: PRT
299 <213> ORGANISM: E. Coli
301 <400> SEQUENCE: 27
302 Met Gly Lys Leu Gly Glu Asn Val Pro Leu Leu Ile Asp Lys Ala Val
303   1             5             10             15
304 Asp Phe Met Ala Ser Ser Gln Ala Phe Arg Glu Tyr Leu Lys Lys Leu
305             20             25             30

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/627,007

DATE: 12/07/2004

TIME: 15:54:37

Input Set : A:\NIH210.001C1 SEQLIST.TXT

Output Set: N:\CRF4\12072004\J627007.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No  
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date